Identification of a novel antimicrobial peptide from the sea star *Patiria pectinifera*

Chan-Hee Kim[a,1], Hye-Jin Go [a,1], Hye Young Oh [a], Ji Been Park [a], Tae Kwan Lee [a], Jung-Kil Seo [b], Maurice R. Elphick[c], and Nam Gyu Park [a,*]

[a]Department of Biotechnology, College of Fisheries Sciences, Pukyong National University, Busan 48513, Korea

[b]Department of Food Science and Biotechnology, Kunsan National University, Kunsan 54150, Korea

[c]School of Biological and Chemical Sciences, Queen Mary University of London, London, E1 4NS UK

*Correspondence to: N. G. Park, Department of Biotechnology, College of Fisheries Sciences, Pukyoung National University, 45 Youngso-ro, Nam-gu, Busan, 48513, Korea; Tel.: +82-51-629-5867; Fax: +82-51-629-5863; *E-mail*: ngpark@pknu.ac.kr (N.G. Park)

Running title: *Pp*CrAMP: a novel sea star cysteine-rich antimicrobial peptide

Footnotes: [1]These authors contributed equally to this work.

Abbreviations: AMPs, antimicrobial peptides; MALDI-TOF MS, matrix assisted laser desorption/ionization time-of-flight mass spectrometry; DTT, 1,4-dithiothreitol; TSB, tryptic soy broth; CFU, colony forming unit; RACE, rapid amplification of cDNA ends; *Pp*CrAMP, *Patiria pectinifera* cysteine-rich antimicrobial peptide; TFA, trifluoroacetic acid; RP, reversed-phase; RT-qPCR, real-time quantitative polymerase chain reaction.

Abstract

Antimicrobial peptides (AMPs) are components of innate immunity found in many forms of life. However, there have been no reports of AMPs in sea star (Phylum Echinodermata). Here we report the isolation and characterization of a novel antimicrobial peptide from the coelomic epithelium extract of the sea star *Patiria pectinifera*. The isolated peptide comprises 38 amino acid residues, is cationic (pI 9.2), has four cysteine residues that form two disulfide bonds (C1-C3 and C2-C4), is amidated at the C-terminus, and is designated *P. pectinifera* cysteine-rich antimicrobial peptide (*Pp*CrAMP). Synthetic *Pp*CrAMP identical to the native peptide exhibited the most potent antimicrobial activity compared to analogs with different disulfide bond configurations. Expression analysis of *Pp*CrAMP precursor transcripts revealed constitutive expression in the coelomic epithelium and tube feet of *P. pectinifera*. Analysis of genomic DNA and cDNA encoding the *Pp*CrAMP precursor protein revealed that an intron splits the coding region of the mature peptide into a positively charged N-terminal domain and a C-terminal domain harboring four cysteine residues and a glycine for C-terminal amidation. No significant homology with other known AMPs was observed, while orthologs of *Pp*CrAMP were found in other echinoderm species. These findings indicate that *Pp*CrAMP is the prototype of a family a novel cysteine-rich AMPs that participate in mechanisms of innate immunity in echinoderms. Furthermore, the discovery of *Pp*CrAMP may lead to the identification of related AMPs in vertebrates and protostome invertebrates.

Keywords: sea star, *Patiria pectinifera*, cysteine-rich antimicrobial peptide, innate immunity, echinoderm

1. Introduction

Antimicrobial peptides (AMPs) are evolutionarily ancient molecules produced by a wide variety of organisms (Ganz, 2003; Zasloff, 2002). AMPs are a major component of the immune defense system in invertebrates, which lack a vertebrate-type adaptive immune system (Bulet et al., 2004; Sperstad et al., 2011). Although AMPs exhibit structural diversity, they are commonly defined as being short (10-50 amino acids, AAs) with a net positive charge (+2 to +9) and have been classified into three major groups: (i) linear peptides that form amphipathic α-helices, (ii) cysteine-rich peptides containing one or more disulfide bonds and (iii) peptides with an overrepresentation of one or two AAs (Bulet et al., 2004; Hancock and Lehrer, 1998; Wang et al., 2016; Zasloff, 2002). The peptides are derived from larger precursor proteins (prepropeptides) that consist of a signal peptide, a prosequence, and a mature peptide (Bulet et al., 2004; Liu and Ganz, 1995; Valore and Ganz, 1992). AMPs are not only characterized by direct antibiotic, antifungal, and antiviral activity against a variety of microorganisms but are also involved indirectly in modulation of the innate immunity, including induction of chemokine production and regulation of apoptosis, angiogenesis, and wound healing (Bowdish et al., 2005; Ganz, 2003; Guilhelmelli et al., 2013; Hancock and Sahl, 2006; Oppenheim and Yang, 2005). Because of the development of antibiotic resistance by microorganisms, AMPs have attracted considerable attention in recent years as potential anti-infective therapeutic candidates for the design of new antimicrobial agents (Craik et al., 2013; Gordon et al., 2005; Parachin and Franco, 2014). In this context, isolation of new AMPs is of interest in providing general insights into AMP structure and activity.

Marine organisms live in habitats abundant with bacteria, fungi, viruses, and parasites, some of which are potentially harmful. However, many marine organisms do not seem to be vulnerable to pathogenic invasions, suggesting that they have robust and effective immune effectors such as AMPs to defend against microbial pathogens (Cheung et al., 2015; Falanga et al., 2016; Otero-Gonzalez et al., 2010). Furthermore, AMPs from marine organisms are often taxon-specific or even species-specific and are structurally different from their counterparts produced by terrestrial species (Augustin et al., 2009; Charlet et al., 1996; Lee et al., 1997; Li et al., 2010b; Li et al., 2008; Smith et al., 2008). Therefore, marine organisms provide fascinating sources for biochemical isolation of novel AMPs.

Echinoderms are a phylum of exclusively marine invertebrates that include sea star, sea urchins, sand dollars, sea cucumbers, and sea lilies. As deuterostome invertebrates, they occupy an intermediate phylogenetic position with respect to the vertebrates and protostome invertebrates and therefore they are of particular interest from an evolutionary perspective (Blair and Hedges, 2005; Smith et al., 2010). Echinoderms rely on innate immunity for defense against harmful microorganisms

and although they are the second largest deuterostome phylum, relatively few AMPs have been isolated and characterized from these animals. Cysteine-rich AMPs (strongylocins) isolated from the sea urchins *Strongylocentrotus droebachiensis, Strongylocentrotus purpuratus*, and *Echinus esculentus* (Li et al., 2010a; Li et al., 2008; Solstad et al., 2016) and heterodimeric AMPs (centrocins) isolated from *S. purpuratus* and *E. esculentus* (Li et al., 2010b; Solstad et al., 2016) exhibit antimicrobial activity against both gram-positive and gram-negative bacteria (Li et al., 2010a; Li et al., 2010b). Strongylocins and centrocins have unique structural characteristics compared to other known AMPs and therefore it is of interest to identify AMPs in other echinoderms (e.g. sea star).

Here we report the isolation of a novel sea star cysteine-rich AMP, named *Pp*CrAMP, from the sea star *Patiria pectinifera*. The primary structure of *Pp*CrAMP was determined by Edman degradation and MALDI-TOF MS and the cysteine connectivity of four cysteine residues that form two disulfide bonds in *Pp*CrAMP was determined by comparison of native and synthetic peptides that were produced with different combinations of two disulfide bond pairings. The antimicrobial activity of synthetic *Pp*CrAMP variants was showed both gram-positive and gram-negative bacteria. Genomic DNA and cDNA encoding the *Pp*CrAMP precursor protein were cloned and sequenced, enabling investigation of its expression pattern in *P. pectinifera*. Furthermore, the organization of the *PpCrAMP* gene in *P. pectinifera* was compared with homologs in other echinoderms. Discovery of *Pp*CrAMP is notable as it is the first cysteine-rich AMP to be purified from sea star.

2. Materials and Methods

2.1. Animals and sample collection

Specimens of the sea star *Patiria pectinifera* were collected at low tide from the intertidal zone on the rocky coast of Cheongsapo of Busan, Korea. The sea star were immediately transferred to our laboratory and maintained in a recirculating seawater system at 15 °C until sample collection. The coelomic epithelium, which includes layers of longitudinal and circular muscle, was collected from the aboral body wall of the arms of 100 specimens of *P. pectinifera* using sterile knives and forceps. The collected sample were immediately frozen in liquid nitrogen and stored at -80 °C until extraction. For immune challenge experiments, 30 live specimens of the sea star (approximate size 4-5 cm determined by the distance from the center of disk to outer tip of arm) were acclimatized in a 600 L recirculating aquarium tank equipped with sand-filtered and UV-sterilized seawater at 15 °C for 1 month. The sea star were fed once every 3 days with live manila clam, *Ruditapes philippinarum*. Approval by the local institution/ethics committee was not required for this work because experimental work on sea star is not subject to regulation and *P. pectinifera* is not an endangered or protected species.

2.2. Peptide extraction and purification

Four volumes of 1% acetic acid was added to the frozen sample and then the mixture was heated in a double boiler for 5 min to prevent proteolytic enzyme activity. The boiled sample was cooled on ice and homogenized (T10 basic ULTRA-TURRAX Homogenizer system, IKA, USA). The homogenate was then centrifuged (20,000 × g, 30 min, 4 °C) and then the supernatant was applied onto a C18 cartridge (Sep-pak C18, 20 cc, Waters Corp, USA). The column was washed with 40 ml of 10% methanol/0.1% trifluoroacetic acid (TFA) and retained materials were then eluted with 40 ml of 60% methanol/0.1% TFA. An aliquot of the eluate was lyophilized and then dissolved in 0.01% acetic acid to evaluate its antimicrobial activity against *Escherichia coli* D31 and *Bacillus subtilis* KCTC1021. To purify antimicrobial components of the eluate, a portion (3 ml) of it was applied to a cation-exchange column (TSKgel SP-5PW, 7.5 × 75 mm, Tosho, Japan) and eluted with a linear gradient of 0 to 1.0 M sodium chloride in 10 mM phosphate buffer (PB, pH 6.0) for 100 min at a flow rate of 1.0 ml/min. Absorbance peaks were monitored at 220 nm to detect peptide bonds and fractions were collected manually. A bioactive peak from the first cation-exchange HPLC purification was subjected to reversed phase (RP)-HPLC (Capcellpak C18, 5μm, 4.6 × 250 mm; Shisheido Co., Tokyo, Japan). Elution was performed by isocratic elution in 10% acetonitrile/0.1% TFA for 10 min and then a linear gradient of 10 to 60% acetonitrile/0.1% TFA for 50 min at a flow rate of 1.0 ml/min. An active peak showing antimicrobial activity against *B. subtilis* KCTC 1021 was purified by

chromatography again using the same column as the previous step, but with isocratic elution in 22% acetonitrile/0.1% TFA (peak A) at a flow rate of 1 ml/min.

2.3. Primary structure determination of the purified peptide

The molecular mass and AA sequence of the purified AMP were determined using matrix assisted laser desorption/ionization (MALDI) time-of-flight (TOF) mass spectrometry (MS) with a pulsed smartbeam II (355 nm Nd:YAG laser, repetition rate 1 kHz) in linear mode (Ultraflextreme from Bruker Daltonics, Billerica, MA, USA) and an automated N-terminal AA gas-phase sequencer (PPSQ-31A/33A protein sequencers, Shimadzu Co., Kyoto, Japan). To confirm the existence of disulfide bonds, the purified peptide was reduced with 100 μl of 0.1 M 1,4-dithiothreitol (DTT) solution for 2 h at 42 °C. After reduction of disulfide bonds, the retention times of the reduced and the native peptides were compared using RP-HPLC with a linear gradient of 5 - 65% acetonitrile/0.1% TFA for 60 min at a flow rate of 1 ml/min.

2.4. Peptide synthesis and determination of disulfide bridge connectivity

Based on the results of structure analyses and cDNA cloning, variants of *Pp*CrAMP with three possible disulfide bond connectivities were custom synthesized by ChemPep Inc. (Wellington, FL, USA). The synthetic peptides were re-purified to be greater than 98% pure by RP-HPLC, and molecular masses were confirmed by MALDI-TOF MS. The reduced form of the synthetic peptide was obtained by RP-HPLC purification followed by the same procedure used for native *Pp*CrAMP. The molecular mass of the reduced synthetic *Pp*CrAMP was also confirmed by MALDI-TOF MS with observation of a 4 mass unit difference. Identity was assessed by comparison of the retention times of synthetic peptides and native *Pp*CrAMP using RP-HPLC with a linear gradient of 20 to 30% acetonitrile/0.1% TFA for 20 min and, then, an isocratic elution with 23% acetonitrile/0.1% TFA. The quantities of the purified synthetic peptides were calculated using a linear relationship between peak area and peptide amount in a serial dilution of 1 mg/ml of synthetic *Pp*CrAMP.

2.5. Antimicrobial activity assay

An ultrasensitive radial diffusion assay was adopted for monitoring antimicrobial activity during the purification steps and for testing synthetic peptides, as described previously (Seo et al., 2016). The microbial strains used to evaluate the antimicrobial activity were *B. subtilis* KCTC1021, *Staphylococcus aureus* KCTC1621, *Micrococcus luteus* KCTC1071, *E. coli* D31, *Streptococcus iniae* FP5229, *Salmonella enterica* ATCC13311, *Shigella flexneri* KCTC2517, *Aeromonas hydrophila* KCTC2358, *Edwardsiella tarda* NUF251, and *Vibrio parahaemolyticus* KCCM41664, and *Candida albicans* KCTC9765 (Table 1). Briefly, microbial strains were pre-cultured overnight in tryptic soy

broth (TSB) at the appropriate temperatures, 25 °C for fish pathogens and 37 °C for the others. Pre-cultured microbial strains were diluted with 10 mM PB (pH 6.6) to ~10^8 CFU/ml for microbial strains and ~10^6 CFU/ml for the fungus *C. albicans*, and 0.5 ml of the diluted strains was mixed with 9.5 ml of underlay gel containing 0.03% TSB and 1% Type I agarose in10 mM PB (pH 6.6). Peptides were serially diluted 2-fold in 5 μl of 0.01% acetic acid and each dilution was added to 2.5 mm diameter wells made in the 1 mm thick underlay gels. After 3 h of incubation at the appropriate temperatures, microbial strains were overlaid with 10 ml of double-strength overlay gel containing 6% TSB with 10 mM PB (pH 6.6) in 1% agarose. Plates were incubated for an additional 18 - 24 h and then the clear zone diameters were measured. After subtracting the diameter of the well, the clear zone diameter was expressed in units (0.1 mm = 1 U). The minimal effective concentration (MEC, μg/ml) of the synthetic peptides was calculated as the X-intercept of a plot of units against the $\log_{10}$ of the peptide concentration (Lehrer et al., 1991). The antimicrobial assay was performed in triplicate and the results were averaged.

2.6. cDNA and gene cloning

Cloning of a cDNA encoding the complete *Pp*CrAMP precursor protein was performed by 3' and 5' rapid amplification of cDNA ends (RACE) polymerase chain reaction (PCR). Total RNA was extracted from the coelomic epithelium of *P. pectinifera* using Hybrid-R kit (GeneAll, Seoul, Korea), and then mRNA was purified using Oligotex mRNA mini kit (Qiagen, USA) following the manufacturer's instructions. The synthesis for RACE-ready cDNA template was performed with GeneRacer kit (RLM-RACE, Invitrogen, CA, USA) according to the manufacturer's instructions. The sequence of primers for 3'RACE was based on analysis of a GenBank transcriptome shotgun assembly (TSA) database (accession no. GFOQ01277783.1) from *P. pectinifera* obtained by Illumina HiSeq 2500 sequencing, reported previously by our group (Kim et al., 2017). Two sequence specific primers were designed for 3' RACE, and then 5' RACE was conducted with sequence-specific primers designed from the sequencing result of the 3' RACE product. The sequences of primers used in RACE are listed in Table 1. The first 3'RACE reaction (30 cycles, 95 °C for 30 s, 60°C for 30 s, and 72 °C for 1 min) was performed using a primer (GSP-F1) and the GeneRacer 3′ primer. The PCR product was re-amplified (30 cycles, 95 °C for 30 s, 58°C for 30 s, and 72 °C for 1 min) using a primer (GSP-F2) and GeneRacer 3′ nested primer. The 5' RACE reaction (30 cycles, 95 °C for 30 s, 58°C for 30 s, and 72 °C for 1 min) was completed using a gene-specific primer (GSP-R) and the GeneRacer 5′ primer. PCR products in the last step of 3' and 5' RACE were introduced into the pGEM-Teasy vector system (Promega Corporation, USA) and sequenced. The sequence of precursor transcripts obtained was submitted to the GenBank database (accession no. MF443207).

Based on the cDNA sequence of the *Pp*CrAMP precursor, both forward (Gene F) and reverse (Gene R) primers located in the 5' and 3' untranslated regions (UTRs) of the cDNA sequence were designed for studying the gene structure (see Table S1 for sequences). Genomic DNA was extracted from the coelomic epithelium of one animal using Exgene DNA extraction kit (GeneAll, Seoul, Korea) following the manufacturer's instructions and 100 ng of genomic DNA was employed as a template in PCR (30 cycles, 95 °C for 30 s, 58°C for 30 s, and 72 °C for 3 min). The PCR product was also cloned into pGEM-Teasy vector and sequenced. The sequence of genomic DNA containing the *PpCrAMP* gene was also submitted to the GenBank database (accession no. MF443208).

2.7. Real time quantitative polymerase chain reaction (RT-qPCR) of PpCrAMP precursor transcripts

RT-qPCR analysis was done to determine the basal expression level of *PpCrAMP* precursor transcripts in various tissues, including coelomic epithelium, coelomocytes, gonad, oral hemal ring including Tiedemann's bodies, pyloric caeca, stomach (including cardiac and pyloric regions), and tube feet. Furthermore, to determine whether acute changes in the abundance of *PpCrAMP* precursor transcripts occur following immune stimulation, tissues that express *PpCrAMP* constitutively (coelomic epithelium and tube feet) were sampled 0, 8, 16, and 32 h post-immune stimulation. The immune challenge was performed by injection with 50 μl *V. parahaemolyticus* (OD_{600}=0.1, 3.3×10^8 CFU/ml) into the coelomic cavity through the body wall at the tip of each of the arms of sea star with arm lengths of 4-5 cm. Total RNA was extracted from pooled sample tissues (five individuals per pool) using Hybrid-R (GeneAll, Seoul, Korea) according to the manufacturer's instructions, and RNA quality was assessed by 1.0% agarose gel electrophoresis and then quantified spectrophotometrically using a NanoDrop Lite (Thermo Fisher Scientific, Wilmington, MA, USA). cDNA was synthesized using the TOPscript cDNA synthesis Kit with oligo dT (dT18) (Enzynomics, Deajeon, Korea) according to the manufacturer's instructions. The primer pairs used for amplifying *PpCrAMP* precursor cDNA and *elongation factor 1α* (*EF1α*, accession No. AAT06175) cDNA as a control for normalization were *PpCrAM*P qPCR-F and qPCR-R, and *EF1α* qPCR-F and qPCR-R, respectively (see Table S1 for sequences). To analyze expression of *PpCrAMP* precursor transcripts in different sea star tissues/organs quantitatively, RT-qPCR was employed using a CFX Connect Real-Time PCR Detection System (Bio-Rad, USA), as previously described with slight modifications (Kim et al., 2016). In brief, the amplification was carried out in a 20 μl reaction mixture containing 10 μl of 2× SYBR green premix (TOPreal qPCR 2X PreMix, Enzynomics, Deajeon, Korea), 1 μl (10 pmol/μl) each of forward and reverse primers, 1 μl of 10 times diluted cDNA template and nuclease free water. The thermal profile was 95 °C for 10 min, 40 cycles of 95 °C for 10 s, 60 °C for 15 s and 72 °C for 15 s with fluorescence recording at the end of each cycle. Melt curve analysis was performed to ensure

product specificity over the temperature range of 60-90 °C. Amplicons were analyzed on agarose gels to confirm the product size. Based on the standard curves for both *PpCrAMP* and *EF1α*, the relative expression levels of *PpCrAMP* precursor transcripts in each tissue were normalized against the level of the *EF1α* control using the comparative CT method ($2^{-\Delta\Delta CT}$) (Livak and Schmittgen, 2001). Triplicate amplifications were carried out independently, and the results were analyzed statistically. For statistical analysis of *PpCrAMP* precursor transcript expression, the graphs were generated, and one-way analysis of variance (ANOVA) with Duncan's multiple range post-hoc analysis was performed using GraphPad Prism software version 7.0 for Windows (GraphPad Software, San Diego, California, USA). Relative fold expression was presented as means ± standard deviation. P values with $p < 0.05$ were considered statistically significant.

2.8. In silico analysis

A cDNA encoding the *Pp*CrAMP precursor protein was translated into protein sequence using Expert Protein Analysis System (ExPASy) proteomics server of the Swiss Institute of Bioinfromatics (http://web.expasy.org/translate/) and SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/) was used to predict the signal peptide of the translated protein sequence. Theoretical molecular mass and isoelectric points of the mature *Pp*CrAMP were calculated by the computer pI/Mw tools (http://web.expasy.org/compute_pi/) at ExPASy. To identify potential homologs of *Pp*CrAMP the deduced AA sequence and genomic nucleotide sequence of the *Pp*CrAMP were submitted as queries for BLAST analysis of i) the NCBI/GenBank nr database (http://blast.ncbi.nlm.nih.gov/blast.cgi), ii) the AMP database, including the collection of antimicrobial peptides (CAMP, http://www.camp.bicnirrh.res.in), iii) the antimicrobial peptide database (APD, http://aps.unmc.edu/AP/main.php), iv) the Echinoderm genomic database (http://www.echinobase.org/Echinobase/Blasts) and v) neural transcriptome sequence data from the sea star *Asterias rubens* (Semmens et al., 2013; Semmens et al., 2016). Multiple sequence alignment of the full-length *P. pectinifera Pp*CrAMP precursor and putative related proteins from other species was performed using a multiple sequence alignment algorithm, Kalign, from the European Bioinformatics Institute (EMBL-EBI) (https://www.ebi.ac.uk/Tools/msa/kalign/). Secondary structure prediction was performed using the Network Protein Sequence Analysis (NPS@: https://npsa-prabi.ibcp.fr/) server (Combet et al., 2000).

3. Results

3.1. Purification of AMP from the coelomic epithelium of Patiria pectinifera

An aliquot of a coelomic epithelium extract of *P. pectinifera* exhibited antimicrobial activity against *B. subtilis* and *E. coli*, which was abolished by tryptic digestion (Fig. 1A), indicating that it was an appropriate source to isolate AMPs. The gram-positive bacterium *B. subtilis* was highly susceptible to the crude extract and so was used to test for antimicrobial activity during the purification steps. A single absorbance peak (peak A) that exhibited antimicrobial activity against *B. subtilis* was purified successfully from the coelomic epithelium extract through three steps of column purification. The extract was first fractionated using cation-exchange HPLC with a salt gradient and an active peak was eluted with 0.6 M sodium chloride corresponding to a retention time of 69 min (Fig. 1B). The peak was further subjected to RP-HPLC and an active peak, designated as peak A, was eluted with 22% acetonitrile/0.1% TFA (Fig. 1C). Finally, a single absorbance peak was obtained with isocratic 22% acetonitrile/0.1% TFA elution, and this peak was then subjected to structural analyses (Fig. 1D).

3.2. Primary structure analyses of purified AMP

The first 37 AAs from the N-terminus of the purified peptide were determined by Edman sequencing (Fig. 2A), but with some unidentified residues (X) from blank cycles. The molecular mass determined by MALDI-TOF MS was 4027.8 Da and 2014.7 as the protonated molecular ion $(M+H)^+$ and the double charged ion $(M+2H)^{2+}$, respectively (Fig. 2B upper panel). Without reduction and alkylation cysteine residues often emerge as blank cycles during amino acid sequencing because they form disulfide bonds that are important for the folding and stability of AMPs and proteins. Therefore, the purified AMP from *P. pectinifera* was reduced by treatment with DTT to confirm the existence of disulfide bonds. The retention time of the reduced peptide was revealed as 29.8 min, which represented a delay of about 2 min compared to the native form (Fig 2C), and the molecular mass of the reduced peptide was 4 Da higher than the native peptide (Fig. 2B lower panel). These data indicated that the purified native peptide contained four cysteine residues that formed two intramolecular disulfide bonds. Accordingly, in the deduced sequence of the AMP we replaced three X residues with cysteine residues and added an additional cysteine residue at the C-terminus: GRKGRKGVRGNPFFNCEDEFGNPGCVCDKRKGGAAVTC. This peptide was designated *P. pectinifera* cysteine-rich antimicrobial peptide (*Pp*CrAMP). The theoretical molecular mass of the deduced peptide in reduced form was calculated as 4032.6 Da $(M+H)^+$, which differed from observed molecular mass of reduced *Pp*CrAMP by 1 Da (Fig. 2B lower panel). C-terminal amidation is a common post-translational modification of AMPs and this decreases the molecular mass by only 1 Da compared to peptides with a free carboxyl-terminus. Furthermore, glycine is a substrate for C-terminal amidation. To investigate if *Pp*CrAMP was C-terminally amidated in this way, the AA sequence of

*Pp*CrAMP was submitted as a query against the non-redundant protein sequences in the NCBI database using BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) but no significant sequence homology with other known AMPs was observed. Therefore, we attempted to find putative transcripts encoding *Pp*CrAMP in GenBank transcriptome shotgun assemblies (TSA) of *Patiria* (taxid: 35076) using BLAST. Two transcripts were found that encoded proteins identical or similar to the AA sequence of *Pp*CrAMP: a 1,242 bp transcribed RNA (accession No. GFOQ01277783.1) and a 910 bp transcribed RNA (accession No. GAWB01039446.1) from *de novo* assembled transcriptomes of *P. pectinifera* and *P. miniata*, respectively (Fig. 2D). The transcripts encoded *Pp*CrAMP or a *Pp*CrAMP-like protein with a glycine residue at its C-terminus, consistent with this residue being a substrate for amidation mediated by peptidylglycine α-amidating monooxygenase (PAM) (Eipper et al., 1991) and the mature *Pp*CrAMP peptide having an α-amide at the C-terminus. In conclusion, the structural analyses demonstrated that *Pp*CrAMP was a C-terminally amidated cationic AMP (with a predicted isoelectric point (pI) of 9.20; http://web.expasy.org/compute_pi/) comprising 38 AAs, which include four cysteine residues (Cys^{16}, Cys^{25}, Cys^{27} and Cys^{38}) that form two disulfide bonds (Fig. 2D).

3.3. Determination of cysteine connectivity in native PpCrAMP

The four cysteine residues in *Pp*CrAMP could mediate three different cysteine connectivities to form two intramolecular disulfide bonds. To determine the authentic cysteine connectivity in the native peptide, we synthesized C-terminally amidated *Pp*CrAMPs that have the three different cysteine connectivities: *Pp*CrAMP-1 (Cys^{16}-Cys^{25} and Cys^{27}-Cys^{38}), *Pp*CrAMP-2 (Cys^{16}-Cys^{27} and Cys^{25}-Cys^{38}) and *Pp*CrAMP-3 (Cys^{16}-Cys^{38} and Cys^{25}-Cys^{27}) corresponding to C1-C2 and C3-C4, C1-C3 and C2-C4, and C1-C4 and C2-C3, respectively (Fig. 3A). The retention time of native *Pp*CrAMP was compared with the elution times of the synthetic peptides using RP-HPLC. Native *Pp*CrAMP was eluted at 15.1 min in a gradient elution, which was almost identical to the retention time (14.9 min) of synthetic *Pp*CrAMP-2. In contrast, both synthetic *Pp*CrAMP-1 and *Pp*CrAMP-3 and the reduced form of *Pp*CrAMP ($PpCrAMP_{reduced}$) were eluted at 16.6, 16.8 and 19.1 min on the same RP-HPLC, respectively, which represented delays of 2 to 4 min compared to the native peptide (Fig. 3B). Furthermore, native *Pp*CrAMP and synthetic *Pp*CrAMP-2 co-eluted with isocratic RP-HPLC (Fig. 3C), whereas the retention times of both synthetic *Pp*CrAMP-1 and *Pp*CrAMP-3 were not identical to synthetic *Pp*CrAMP-2 (Fig. 3D and E). Collectively, these findings indicated that the four cysteine residues in native *Pp*CrAMP formed two disulfide bonds with Cys^{16}-Cys^{27} and Cys^{25}-Cys^{38} pairings (i.e. C1-C3, C2-C4 connectivity) and the C-terminus of native *Pp*CrAMP was amidated.

3.4. Antimicrobial activity of synthetic PpCrAMP variants

All four synthetic *Pp*CrAMP variants exhibited antimicrobial activity against the gram-negative bacteria *S. enterica* and *S. flexneri*, with a minimal effective concentration (MEC) of 4.5 to 31.4 μg/ml, and against the gram-positive bacteria *B. subtilis*, *S. aureus*, and *M. luteus*, with a MEC of 15.6 to >250 μg/ml. However, the antimicrobial activity of synthetic *Pp*CrAMP-2 was significantly higher than the antimicrobial activity of the other synthetic *Pp*CrAMPs and reduced *Pp*CrAMP (Table 1). The most potent antimicrobial activity exhibited by all four synthetic *Pp*CrAMPs was against the gram-negative bacterium *S. enterica* [MECs, 4.5 – 8.4 μg/ml]. However, antimicrobial activity of synthetic *Pp*CrAMPs was barely detectable against fish pathogens and was undetectable with the fungus *C. albicans*. These findings demonstrated that the existence of the disulfide bonds in *Pp*CrAMP was not critical for antimicrobial activity, but the cysteine connectivity in native *Pp*CrAMP corresponding to C1-C3 and C2-C4 was required for maximum activity against the bacteria tested here. Interestingly, while coelomic epithelium extract showed antimicrobial activity against *E. coli* D31 (Fig. 1A), all four synthetic *Pp*CrAMPs did not show antimicrobial activity against *E. coli* D31 up to a peptide concentration of 250 μg/ml, indicating that the coelomic epithelium extract also contained other AMPs responsible for antimicrobial activity against *E. coli* D31.

3.5. cDNA and genomic DNA sequence encoding PpCrAMP

To obtain the complete sequence of the *Pp*CrAMP precursor protein, a cDNA encoding *Pp*CrAMP was cloned and sequenced (accession number: MF443207). The cDNA of the *Pp*CrAMP precursor comprised 926 bp, starting with a 5'-UTR of 81 bp, followed by an open reading frame (ORF) of 240 bp, a 3'-UTR of 605 bp containing a polyadenylation consensus sequence (AATAAA) located at 31 bp upstream of a poly(A)$^+$ tail (Fig. 4A). The deduced AA sequence of the ORF of the *Pp*CrAMP precursor started with a signal peptide of 21 residues, as predicted by SignalP 4.1, followed by two peptide fragments cleaved at putative dibasic cleavage site ($Lys^{39}Arg^{40}$): an N-terminal prosequence (Ser^{22}-Val^{38}) containing several anionic AAs and mature *Pp*CrAMP consisting of 38 AAs plus one C-terminal glycine residue (Gly^{41}-Gly^{79}), consistent with the structural analyses (Fig. 4B). Accordingly, these sequence data demonstrated that *Pp*CrAMP was derived from a larger precursor protein which underwent post-translational modifications such as formation of disulfide bonds and α-amidation at the C-terminus followed by cleavage at a putative dibasic cleavage site (KR) between the anionic prosequence and the mature peptide. The genomic DNA sequence encoding *Pp*CrAMP (accession number: MF443208) comprised two exons and one intron (Fig. 4B). The first exon comprised a 5′ UTR followed by an ORF encoding the signal peptide, the prosequence, and the first 14 AAs of mature *Pp*CrAMP, which was followed by an 896 bp intron. The second exon comprised an ORF encoding the cysteine-rich region of *Pp*CrAMP (25 AAs) followed by a 3′ UTR.

The classical canonical splicing recognition sequence GT/AG was present at the exon-intron junctions.

BLAST analysis revealed that *Pp*CrAMP exhibited no significant sequence homology with other known AMP precursors. However, genomic DNA sequence encoding *Pp*CrAMP exhibited sequence similarity with genes in the sea star *P. miniata* (accession No. AKZP01101613), the sea star *Acanthaster planci* (accession No. BDGH01001773), the sea cucumber *P. parvimensis* (accession No. JXUT0100825), and the sea urchin *S. purpuratus* (accession No. AAGJ05078965) (Fig. 5A). All four genes were similar to the *P. pectinifera Pp*CrAMP gene in containing one intron and two exons, which encoded homologs of the *Pp*CrAMP precursor. In addition, analysis of neural transcriptome sequence data from the sea star *Asterias rubens* (Semmens et al., 2013; Semmens et al., 2016) revealed two transcripts encoding homologs of the *Pp*CrAMP precursor – *Ar*CrAMP-1 precursor (accession number: MG711458) and *Ar*CrAMP-2 precursor (accession number: MG711459). A multiple alignment of the *Pp*CrAMP precursor with homologs identified in other echinoderms is shown in Fig. 5B. The *Pp*CrAMP precursor shared 96.2% AA identity with the homolog from the sea star *P. miniata*, 53.6% AA identity with the homolog from the sea star *A. planci*, 45.7% and 46.6% identity with the two homologs from the sea star *A. rubens*, 43.1% AA identity with the homolog from the sea urchin *S. purpuratus*, and 34.3% AA with the homolog from the sea cucumber *P. parvimensis*. Collectively, these data indicated that *Pp*CrAMP was the prototype for a novel family of cysteine-rich AMPs that occur in echinoderms.

3.6. RT-qPCR analysis for PpCrAMP mRNA

To compare expression levels of *PpCrAMP* transcript in various sea star tissues and post immune challenge, the relative expression levels of the *PpCrAMP* precursor transcript in different tissues (coelomic epithelium, coelomocytes, gonad, oral hemal ring, pyloric caeca, stomach, and tube feet) of *P. pectinifera* were determined by RT-qPCR using sequence specific primers targeting the *PpCrAMP* coding region. An *EF1α* gene was used as an invariant control and for comparison of relative expression between transcripts (Kim et al., 2016). The results showed that the highest expression level of *PpCrAMP* precursor transcripts was detected in the tube feet and the coelomic epithelium, which was the original source of *Pp*CrAMP in this study, followed by moderate expression levels in the oral hemal ring (including Tiedemann's bodies), and the stomach (Fig. 6A). These findings indicated that the coelomic epithelium and the tube feet were the major tissues/organs that produced *PpCrAMP* in *P. pectinifera*. Accordingly, these two tissues were selected to determine whether acute changes in the abundance of *PpCrAMP* precursor transcript occur after bacterial challenge. However, no significant

changes in *PpCrAMP* precursor expression were observed at different times after the bacterial challenge (Fig. 6B).

4. Discussion

Few AMPs have been identified in echinoderms to date. Strongylocins and centrocins were first isolated from coelomocytes of the green sea urchin (*S. droebachiensis*) and related peptides (SpStronylocins 1 and 2) were then discovered and characterized in the purple sea urchin (*S. purpuratus*) and the edible sea urchin (*E. esculentus*) (Li et al., 2010a; Li et al., 2010b; Li et al., 2008; Solstad et al., 2016). These are cationic peptides that exhibit antimicrobial activity against both gram-positive and gram-negative bacteria (Li et al., 2010a; Li et al., 2008). Strongylocins with six cysteine residues forming three intramolecular disulfide bonds show a cysteine arrangement pattern different from any known cysteine-rich AMPs with six cysteine residues and have post-translational modifications such as a brominated tryptophan (Li et al., 2008). Centrocins have a heterodimeric structure, containing a heavy chain (30 AAs) and a light chain (12 AAs), and also have a brominated tryptophan (Li et al., 2010b). Thus, AMPs isolated from echinoderm species, including strongylocins and centrocins, have distinct structures compared to those that have been isolated from vertebrates and protostomes. Here we report the purification from an extract of the coeleomic epithelium of the sea star *P. pectinifera* of a novel AMP designated *Pp*CrAMP, which contains four cysteine residues that form two disulfide bonds and which has a amidated C-terminal cysteine (Fig. 2 and 3).

Cysteine-rich AMPs represent the most diverse and widely distributed family of AMPs in the animal kingdom. Depending on the number of cysteine residues (mostly between 2 to 8) and their paring, cysteine-rich AMPs are classified into three groups: a β-sheet conformation with triple strands, a β-hairpin-like structure, and a mixed α-helix/β-sheet conformation (Bulet et al., 2004). Among these three groups of peptides, AMPs containing four cysteine residues that form two disulfide bonds have been identified in arthropods and pigs (Fig. 7): tachyplesin and polypemusin from the horseshoe crab *Tachypleus tridentatus* and *Limulus polyphemus* (Miyata et al., 1989; Nakamura et al., 1988), respectively, gomesin from the spider *Acanthoscuria gomesiana* (Silva et al., 2000), androctonin from the scorpion *Androctonus australis* (Ehret-Sabatier et al., 1996), and protegrin from porcine leukocytes (Storici and Zanetti, 1993). Moreover, with exception of androctonin, all of these peptides are amidated at the C-terminus and their cysteine connectivity is C1-C4 and C2-C3 (Fahrner et al., 1996; Laederach et al., 2002; Mandard et al., 2002). In contrast, the novel AMP identified here in the sea star *P. pectinifera*, *Pp*CrAMP, has two disulfide bonds with C1-C3 and C2-C4 connectivity.

The antimicrobial activity of synthetic *Pp*CrAMP, with C1-C3 and C2-C4 cysteine connectivity (*Pp*CrAMP-2), is identical to that of the native peptide, and synthetic *Pp*CrAMP-2 exhibits the most potent activity against both gram-positive and gram-negative bacteria compared with other synthetic variants. Investigation of the importance of the disulfide bonds in cysteine-rich antimicrobial peptides

with two disulfide bonds demonstrates that the peptides require the correct disulfide bond configuration to adopt a conformation such as the β-hairpin-like structure and to retain full bioactivity (Laederach et al., 2002; Mani et al., 2005; Muhle and Tam, 2001; Rao, 1999). The β-hairpin-like structure consists of two antiparallel β-strands stabilized by a disulfide bond, linked by a short loop of two to five amino acids (Panteleev et al., 2015). The β-hairpin-like structure that is essential for the activity seen in cysteine-rich AMPs (e.g. tachyplesin-I and protegrin-I) is consistent with the predicted consensus secondary structure of *Pp*CrAMP (Fig. 7). Analysis of the sequence of *Pp*CrAMP using the NPS@ server indicates that *Pp*CrAMP is likely to adopt a β-hairpin-like structure consisting of two extended β-strands (residues 25-28 and 35-37) linked by a random coil region. Therefore, the potent antimicrobial activity of synthetic *Pp*CrAMP-2 may reflect the disulfide bond connectivity that establishes the most stable structure. Further investigation of the relationship between conformation and antimicrobial activity of *Pp*CrAMP will be required to address this issue. Homologs of *Pp*CrAMP identified in other echinoderms also have four cysteine residues in equivalent positions but C-terminal amidation appears not to be a generic characteristic. For example, the two *Pp*CrAMP-type proteins identified in the sea star *A. rubens* do not have C-terminal glycine residue that could provide a substrate for C-terminal amidation.

The *PpCrAMP* gene contains an intron that interrupts the region of the open reading frame encoding the mature *Pp*CrAMP, with one exon encoding the N-terminal domain and another exon encoding the C-terminal domain that contains four cysteine residues. Orthologous genes in other echinoderm species, including the sea curcumber *P. parvimensis* and the sea urchin *S. purpuratus*, have the same intron/exon structure (Fig. 4 and 5). Although the organization of genes encoding cysteine-rich AMPs is very diverse, the peptides are classified in the same structural scaffold group based upon size, cysteine pattern and function, revealing links between the AMPs found in vertebrates and those found in invertebrates (Charlet et al., 1996; Froy, 2005). Nothing is known about the occurrence of *Pp*CrAMP-like proteins in other phyla. However, the occurrence of *Pp*CrAMP-type proteins in echinoderms, a phylum that occupies an “intermediate” position with respect to the deuterostome invertebrates, which include two chordate subphyla that are closely related to vertebrates and the Ambulacraria, and protostome invertebrates, indicates there is a possibility of the presence of orthologous genes and proteins related to defense in deuterostome invertebrates as well as protostomes.

Analysis of the expression of the *PpCrAMP* precursor transcripts in *P. pectinifera* using qPCR reveals that the coelomic epithelium and the tube feet are a major source of *Pp*CrAMP. This is consistent with our recent finding that the coelomic epithelium and the tube feet are grouped in a

tissue/organ cluster with a related biological functions based on an evaluation of differentially expressed genes in *P. pectinifera* using *de novo* transcriptome data (BioProject accession: PRJNA371229) (Kim et al., 2017). The coelomic epithelium is a tissue layer that lines the aboral inner surface of the body wall of sea star. It appears to be a unique tissue with many features of an "ancient multifunctional organogenetic tissue", which is involved not only in common epithelial functions, but also in a range of important biological processes such as wound healing, regeneration, and haematopoiesis (Bossche and Jangoux, 1976; Holm et al., 2008). The absence of change in the expression levels of the *PpCrAMP* precursor transcripts after immune challenge suggests that *Pp*CrAMP may contribute to innate immune defense in an indirect manner. Recent study on the neuropeptide NDA-1, which was secreted in sensory and ganglion of the ectodermal epithelium of the model organism *Hydra* during early development, surprisingly shows antimicrobial activity that may affect microbiome composition on the body surface (Augustin et al., 2017). *Pp*CrAMP may also contribute to endocrine system with antimicrobial activity in a similar manner on the body surface.

In conclusion, *Pp*CrAMP, the cysteine-rich AMP isolated from the coelomic epithelium of the sea star *P. pectinifera*, is the first reported sea star AMP. This study increases our knowledge of AMPs that are involved in the innate immune system of sea star and other echinoderm species and may lead to the discovery of homologs that are involved in immune mechanisms of other animal types. Furthermore, *Pp*CrAMP along with AMPs isolated from other echinoderms may provide a framework for development of novel antimicrobial drugs.

Acknowledgements

The research was supported by the Korea Ministry of Environment (MOE) "Eco-innovation Program (201300030002)"

5. References

Augustin, R., Anton-Erxleben, F., Jungnickel, S., Hemmrich, G., Spudy, B., Podschun, R., Bosch, T.C., 2009. Activity of the novel peptide arminin against multiresistant human pathogens shows the considerable potential of phylogenetically ancient organisms as drug sources. Antimicrob. Agents Chemother. 53, 5245-5250.

Augustin, R., Schröder, K., Murillo Rincón, A.P., Fraune, S., Anton-Erxleben, F., Herbst, E.M., Wittlieb, J., Schwentner, M., Grötzinger, J., Wassenaar, T.M. and Bosch, T.C.G. 2017. A secreted antibacterial neuropeptide shapes the microbiome of Hydra. Nat. Commun. 8, 698.

Blair, J.E., Hedges, S.B., 2005. Molecular phylogeny and divergence times of deuterostome animals. Mol. Biol. Evol. 22, 2275-2284.

Bossche, J.P., Jangoux, M., 1976. Epithelial origin of starfish coelomocytes. Nature 261, 227-228.

Bowdish, D.M., Davidson, D.J., Scott, M.G., Hancock, R.E., 2005. Immunomodulatory activities of small host defense peptides. Antimicrob. Agents Chemother. 49, 1727-1732.

500 Bulet, P., Stocklin, R., Menin, L., 2004. Anti-microbial peptides: from invertebrates to vertebrates.
501 Immunol. Rev. 198, 169-184.
502 Charlet, M., Chernysh, S., Philippe, H., Hetru, C., Hoffmann, J.A., Bulet, P., 1996. Innate immunity.
503 Isolation of several cysteine-rich antimicrobial peptides from the blood of a mollusc, Mytilus
504 edulis. J. Biol. Chem. 271, 21808-21813.
505 Cheung, R.C., Ng, T.B., Wong, J.H., 2015. Marine Peptides: Bioactivities and Applications. Mar.
506 Drugs 13, 4006-4043.
507 Combet, C., Blanchet, C., Geourjon, C., Deleage, G., 2000. NPS@: network protein sequence
508 analysis. Trends Biochem. Sci. 25, 147-150.
509 Craik, D.J., Fairlie, D.P., Liras, S., Price, D., 2013. The future of peptide-based drugs. Chem. Biol.
510 Drug Des. 81, 136-147.
511 Ehret-Sabatier, L., Loew, D., Goyffon, M., Fehlbaum, P., Hoffmann, J.A., van Dorsselaer, A., Bulet,
512 P., 1996. Characterization of novel cysteine-rich antimicrobial peptides from scorpion blood. J.
513 Biol. Chem. 271, 29537-29544.
514 Eipper, B.A., Perkins, S.N., Husten, E.J., Johnson, R.C., Keutmann, H.T., Mains, R.E., 1991.
515 Peptidyl-alpha-hydroxyglycine alpha-amidating lyase. Purification, characterization, and
516 expression. J. Biol. Chem. 266, 7827-7833.
517 Fahrner, R.L., Dieckmann, T., Harwig, S.S., Lehrer, R.I., Eisenberg, D., Feigon, J., 1996. Solution
518 structure of protegrin-1, a broad-spectrum antimicrobial peptide from porcine leukocytes. Chem.
519 Biol. 3, 543-550.
520 Falanga, A., Lombardi, L., Franci, G., Vitiello, M., Iovene, M.R., Morelli, G., Galdiero, M., Galdiero,
521 S., 2016. Marine antimicrobial peptides: nature provides templates for the design of novel
522 compounds against pathogenic bacteria. Int. J. Mol. Sci. 17.
523 Froy, O., 2005. Convergent evolution of invertebrate defensins and nematode antibacterial factors.
524 Trends Microbiol. 13, 314-319.
525 Ganz, T., 2003. The role of antimicrobial peptides in innate immunity. Integr. Comp. Biol. 43, 300-
526 304.
527 Gordon, Y.J., Romanowski, E.G., McDermott, A.M., 2005. A review of antimicrobial peptides and
528 their therapeutic potential as anti-infective drugs. Curr. Eye Res. 30, 505-515.
529 Guilhelmelli, F., Vilela, N., Albuquerque, P., Derengowski Lda, S., Silva-Pereira, I., Kyaw, C.M.,
530 2013. Antibiotic development challenges: the various mechanisms of action of antimicrobial
531 peptides and of bacterial resistance. Front. Microbiol. 4, 353.
532 Hancock, R.E., Lehrer, R., 1998. Cationic peptides: a new source of antibiotics. Trends Biotechnol.
533 16, 82-88.
534 Hancock, R.E., Sahl, H.G., 2006. Antimicrobial and host-defense peptides as new anti-infective
535 therapeutic strategies. Nat. Biotechnol. 24, 1551-1557.
536 Holm, K., Dupont, S., Skold, H., Stenius, A., Thorndyke, M., Hernroth, B., 2008. Induced cell
537 proliferation in putative haematopoietic tissues of the sea star, Asterias rubens (L.). J. Exp. Biol.
538 211, 2551-2558.
539 Kim, C.H., Go, H.J., Oh, H.Y., Jo, Y.H., Elphick, M.R., Park, N.G., 2017. Transcriptomics reveals
540 tissue/organ-specific differences in gene expression in the starfish Patiria pectinifera. Mar
541 Genom. 37, 92-96.
542 Kim, C.H., Kim, E.J., Go, H.J., Oh, H.Y., Lin, M., Elphick, M.R., Park, N.G., 2016. Identification of a
543 novel starfish neuropeptide that acts as a muscle relaxant. J. Neurochem. 137, 33-45.
544 Laederach, A., Andreotti, A.H., Fulton, D.B., 2002. Solution and micelle-bound structures of
545 tachyplesin I and its active aromatic linear derivatives. Biochem. 41, 12359-12368.
546 Lee, I.H., Cho, Y., Lehrer, R.I., 1997. Styelins, broad-spectrum antimicrobial peptides from the
547 solitary tunicate, Styela clava. Comp. Biochem. Physiol. B, Biochem. Mol. Biol. 118, 515-521.
548 Lehrer, R.I., Rosenman, M., Harwig, S.S., Jackson, R., Eisenhauer, P., 1991. Ultrasensitive assays for
549 endogenous antimicrobial polypeptides. J. Immunol. Methods 137, 167-173.

Li, C., Blencke, H.-M., Smith, L.C., Karp, M.T., Stensvåg, K., 2010a. Two recombinant peptides, SpStrongylocins 1 and 2, from Strongylocentrotus purpuratus, show antimicrobial activity against Gram-positive and Gram-negative bacteria. Dev. Comp. Immunol. 34, 286-292.

Li, C., Haug, T., Moe, M.K., Styrvold, O.B., Stensvag, K., 2010b. Centrocins: isolation and characterization of novel dimeric antimicrobial peptides from the green sea urchin, Strongylocentrotus droebachiensis. Dev. Comp. Immunol. 34, 959-968.

Li, C., Haug, T., Styrvold, O.B., Jørgensen, T.Ø., Stensvåg, K., 2008. Strongylocins, novel antimicrobial peptides from the green sea urchin, Strongylocentrotus droebachiensis. Dev. Comp. Immunol. 32, 1430-1440.

Liu, L., Ganz, T., 1995. The pro region of human neutrophil defensin contains a motif that is essential for normal subcellular sorting. Blood 85, 1095-1103.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods (San Diego, Calif.) 25, 402-408.

Mandard, N., Bulet, P., Caille, A., Daffre, S., Vovelle, F., 2002. The solution structure of gomesin, an antimicrobial cysteine-rich peptide from the spider. Eur. J. Biochem. 269, 1190-1198.

Mani, R., Waring, A.J., Lehrer, R.I., Hong, M., 2005. Membrane-disruptive abilities of β-hairpin antimicrobial peptides correlate with conformation and activity: A 31P and 1H NMR study. Biochim.Biophys. Acta 1716, 11-18.

Miyata, T., Tokunaga, F., Yoneya, T., Yoshikawa, K., Iwanaga, S., Niwa, M., Takao, T., Shimonishi, Y., 1989. Antimicrobial peptides, isolated from horseshoe crab hemocytes, tachyplesin II, and polyphemusins I and II: chemical structures and biological activity. J. Biochem. 106, 663-668.

Muhle, S.A., Tam, J.P., 2001. Design of Gram-negative selective antimicrobial peptides. Biochem. 40, 5777-5785.

Nakamura, T., Furunaka, H., Miyata, T., Tokunaga, F., Muta, T., Iwanaga, S., Niwa, M., Takao, T., Shimonishi, Y., 1988. Tachyplesin, a class of antimicrobial peptide from the hemocytes of the horseshoe crab (Tachypleus tridentatus). Isolation and chemical structure. J. Biol. Chem. 263, 16709-16713.

Oppenheim, J.J., Yang, D., 2005. Alarmins: chemotactic activators of immune responses. Curr. Opin. Immunol. 17, 359-365.

Otero-Gonzalez, A.J., Magalhaes, B.S., Garcia-Villarino, M., Lopez-Abarrategui, C., Sousa, D.A., Dias, S.C., Franco, O.L., 2010. Antimicrobial peptides from marine invertebrates as a new frontier for microbial infection control. FASEB J. 24, 1320-1334.

Panteleev, P.V., Bolosov, I.A., Balandin, S.V., Ovchinnikova, T.V., 2015. Structure and biological functions of β-hairpin antimicrobial peptides. Acta Naturae 7, 37-47.

Parachin, N.S., Franco, O.L., 2014. New edge of antibiotic development: antimicrobial peptides and corresponding resistance. Front. Microbiol. 5, 147.

Rao, A.G., 1999. Conformation and antimicrobial activity of linear derivatives of tachyplesin lacking disulfide bonds. Arch. Biochem. Biophys. 361, 127-134.

Semmens, D.C., Dane, R.E., Pancholi, M.R., Slade, S.E., Scrivens, J.H., Elphick, M.R., 2013. Discovery of a novel neurophysin-associated neuropeptide that triggers cardiac stomach contraction and retraction in starfish. J. Exp. Biol. 216, 4047-4053.

Semmens, D.C., Mirabeau, O., Moghul, I., Pancholi, M.R., Wurm, Y., Elphick, M.R., 2016. Transcriptomic identification of starfish neuropeptide precursors yields new insights into neuropeptide evolution. Open Biol. 6, 150224.

Seo, J.K., Go, H.J., Kim, C.H., Nam, B.H., Park, N.G., 2016. Antimicrobial peptide, hdMolluscidin, purified from the gill of the abalone, Haliotis discus. Fish Shellfish Immunol. 52, 289-297.

Silva, P.I., Jr., Daffre, S., Bulet, P., 2000. Isolation and characterization of gomesin, an 18-residue cysteine-rich defense peptide from the spider Acanthoscurria gomesiana hemocytes with sequence similarities to horseshoe crab antimicrobial peptides of the tachyplesin family. J. Biol. Chem. 275, 33464-33470.

601 Smith, L.C., Ghosh, J., Buckley, K.M., Clow, L.A., Dheilly, N.M., Haug, T., Henson, J.H., Li, C.,
602 Lun, C.M., Majeske, A.J., Matranga, V., Nair, S.V., Rast, J.P., Raftos, D.A., Roth, M., Sacchi, S.,
603 Schrankel, C.S., Stensvag, K., 2010. Echinoderm immunity. Adv. Exp. Med. Biol. 708, 260-301.
604 Smith, V.J., Fernandes, J.M., Kemp, G.D., Hauton, C., 2008. Crustins: enigmatic WAP domain-
605 containing antibacterial proteins from crustaceans. Dev. Comp. Immunol. 32, 758-772.
606 Solstad, R.G., Li, C., Isaksson, J., Johansen, J., Svenson, J., Stensvag, K., Haug, T., 2016. Novel
607 antimicrobial peptides EeCentrocins 1, 2 and EeStrongylocin 2 from the edible sea urchin
608 Echinus esculentus have 6-Br-Trp post-translational modifications. PLoS ONE 11, e0151820.
609 Sperstad, S.V., Haug, T., Blencke, H.M., Styrvold, O.B., Li, C., Stensvag, K., 2011. Antimicrobial
610 peptides from marine invertebrates: challenges and perspectives in marine antimicrobial peptide
611 discovery. Biotechnol. Adv. 29, 519-530.
612 Storici, P., Zanetti, M., 1993. A novel cDNA sequence encoding a pig leukocyte antimicrobial peptide
613 with a cathelin-like pro-sequence. Biochem. Biophys. Res. Commun. 196, 1363-1368.
614 Valore, E.V., Ganz, T., 1992. Posttranslational processing of defensins in immature human myeloid
615 cells. Blood 79, 1538-1544.
616 Wang, G., Li, X., Wang, Z., 2016. APD3: the antimicrobial peptide database as a tool for research and
617 education. Nucleic Acids Res. 44, D1087-1093.
618 Zasloff, M., 2002. Antimicrobial peptides of multicellular organisms. Nature 415, 389-395.

Figure legends

Fig. 1. Isolation of an AMP from an extract of coelomic epithelium of the sea star *P. pectinifera*. (A) Antimicrobial activity of the crude and the trypsin treated extract against *B. subtilis* KCTC 1021 and *E. coli* D31 is shown. (B) Fractionation of the crude extract by cation-exchange HPLC reveals an active peak (downward arrow) is eluted with 0.6 M sodium chloride. (C). A single absorbance peak (peak A) responsible for the antimicrobial activity against *B. subtilis* was obtained in the second RP-HPLC step. (D) Peak A was isolated through RP-HPLC with isocratic elution in 22% acetonitrile/0.1% TFA.

Fig. 2. Primary structure determination of purified AMP. (A) N-terminal 37 amino acid residues of purified peak A was determined by Edman degradation. (B) The retention times of the native peptide and the reduced peptide (after treatment 0.1 M DTT) on RP-HPLC were compared. (C). MALDI-TOF MS analysis of the native peptide and the reduced peptide showed a 4 Da difference in molecular mass consistent with the presence of two disulfide bonds. (D). Complete primary structure of purified peak A designated *P. pectinifera* cysteine-rich AMP (*Pp*CrAMP) comprised 38 AAs with C-terminal α-amidation and was compared with sequences derived from transcriptome data obtained from *P. pectinifera* and *P. miniata*.

Fig. 3. Determination of the disulfide bond cysteine connectivity of native *Pp*CrAMP (A) Structures of four *Pp*CrAMP variants with three different disulfide bond connectivities or without disulfide bonds are shown. (B) The retention times of native *Pp*CrAMP and synthetic variants were compared using RP-HPLC with a linear gradient of 20% to 30% acetonitrile/0.1% TFA over 20 min. (C) Native *Pp*CrAMP co-elutes with synthetic *Pp*CrAMP that has a C1-C3 and C2-C4 connectivity using RP-HPLC under isocratic conditions with 23% acetonitrile/0.1% TFA. (D) Synthetic *Pp*CrAMP-1 and *Pp*CrAMP-3 co-elute under the same conditions as in (C). (E) Synthetic *Pp*CrAMP-1 and *Pp*CrAMP-2 do not co-elute under the same conditions as in (C).

Fig. 4. Sequence and structural features of the *P. pectinifera Pp*CrAMP precursor. (A) Schematic showing the structure of a cDNA encoding the *Pp*CrAMP precursor protein is shown. (B) DNA sequence of the gene encoding the *Pp*CrAMP precursor protein, which comprises two exons (upper case) separated by an intron (lowercase) is shown. The canonical splicing recognition sequence GT/AG and the polyadenylation signal site are shadow boxed and underlined, respectively. The amino acid sequence of the precursor is shown below the coding sequence, with the predicted signal peptide, anionic prosequence and purified mature *Pp*CrAMP shown in blue, black, and red, respectively, and a putative dibasic cleavage site (KR) shown in green. A glycine residue that provides a substrate for C-terminal amidation is boxed and the stop codon is indicated with an asterisk. The sequences of the

cDNA and genomic DNA encoding the *Pp*CrAMP precursor are accessible from GenBank under accession numbers MF443207 and MF443208, respectively.

Fig. 5. Comparison of the gene structure and sequences of *Pp*CrAMP-type proteins in echinoderms (A). The structure of the gene encoding *Pp*CrAMP in the sea star *P. pectinifera* with related genes in the sea star *P. miniata* and *A. planci*, the sea cucumber *P. parvimensis*, and the sea urchin *S* are compared. *purpuratus*: (B) Sequence alignment of *P. pectinifera Pp*CrAMP with *Pp*CrAMP-like peptides from other echinoderms.

Fig. 6. Quantitative analysis of basal expression of *Pp*CrAMP precursor transcripts in various organs/tissues (A) and after immune challenge in the coelomic epithelium and the tube feet (B) from *P. pectinifera.* The relative expression levels of *Pp*CrAMP transcripts in each organ/tissue were normalized against the level of the *EF1α* gene as an internal control. Means ± standard deviation (n=3) are shown. Means denoted by the same letter did not differ significantly ($p > 0.05$) while different letters (a, b, c, d) at the top of the bars indicate statistically significant differences ($p < 0.05$) between tissues determined by one-way ANOVA followed by Duncan's Multiple Range test.

Fig. 7. Comparison of amino acid sequence and cysteine array of *P. pectinifera Pp*CrAMP to vertebrate and invertebrate cysteine-rich AMPs that have four cysteine residues forming two disulfide bonds and adopting β-hairpin-like structure. *Pp*CrAMP is compared with (i) tachyplesin-I and (ii) polyphemusin-I from the horseshoe crabs *T. tridentatus* and *L. polyphemus*, respectively (Miyata et al., 1989; Nakamura et al., 1988); (iii) gomesin from the spider *A. gomesiana* (Silva et al., 2000); (iv) androctonin from the scorpion *A. australis* (Ehret-Sabatier et al., 1996); (v) protegrin from from porcine leukocytes (Storici and Zanetti, 1993). Lowercase (a) at the C-terminus of peptides and lowercase (p) at the N-terminus of peptides indicate a C-terminal α-amide and pyroglutamate, respectively. Predicted consensus secondary structure of *Pp*CrAMP using NPS@ server is shown below the amino acid sequence. Lowercase c (orange) and e (blue) indicate random coil and extended strand, respectively.

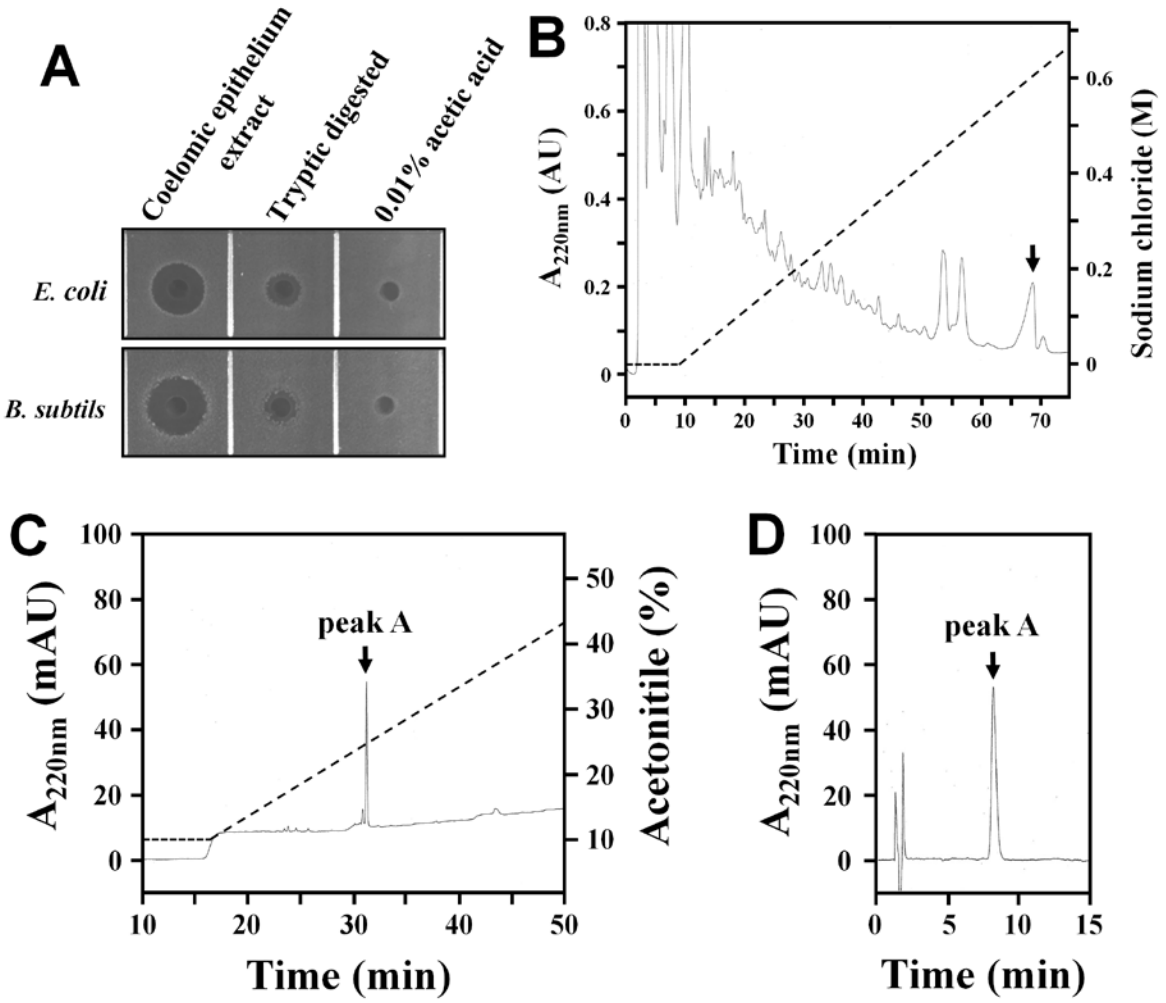
678 **Table legend**

679 **Table 1.** Antimicrobial activity against various microbial strains of synthetic *Pp*CrAMPs, including
680 the reduced peptide and peptides with three different combinations of two disulfide bonds

Supplementary table legend

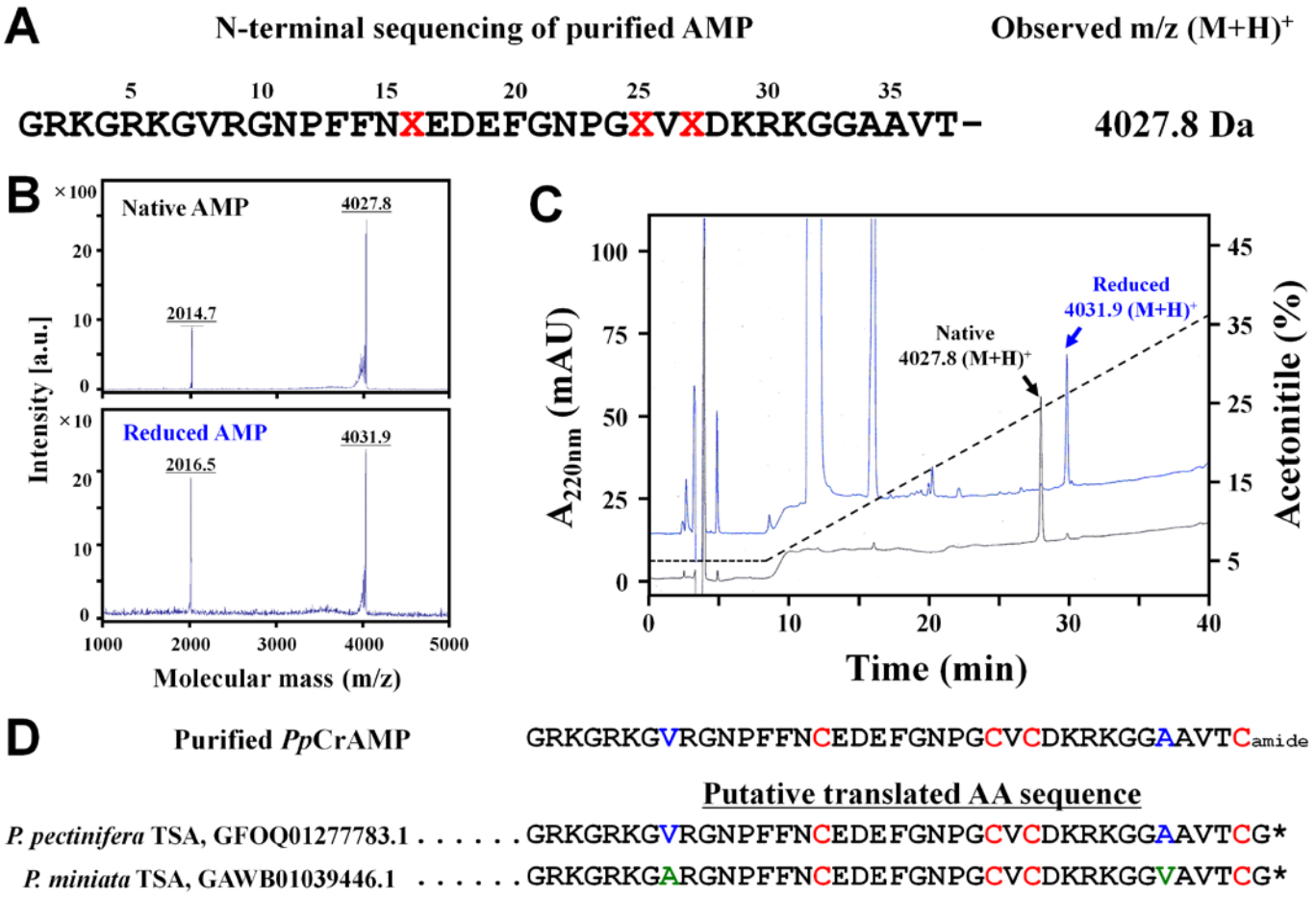
Supplementary Table S1. Designations and nucleotide sequences of the primers used in this study

683 Figure 1.



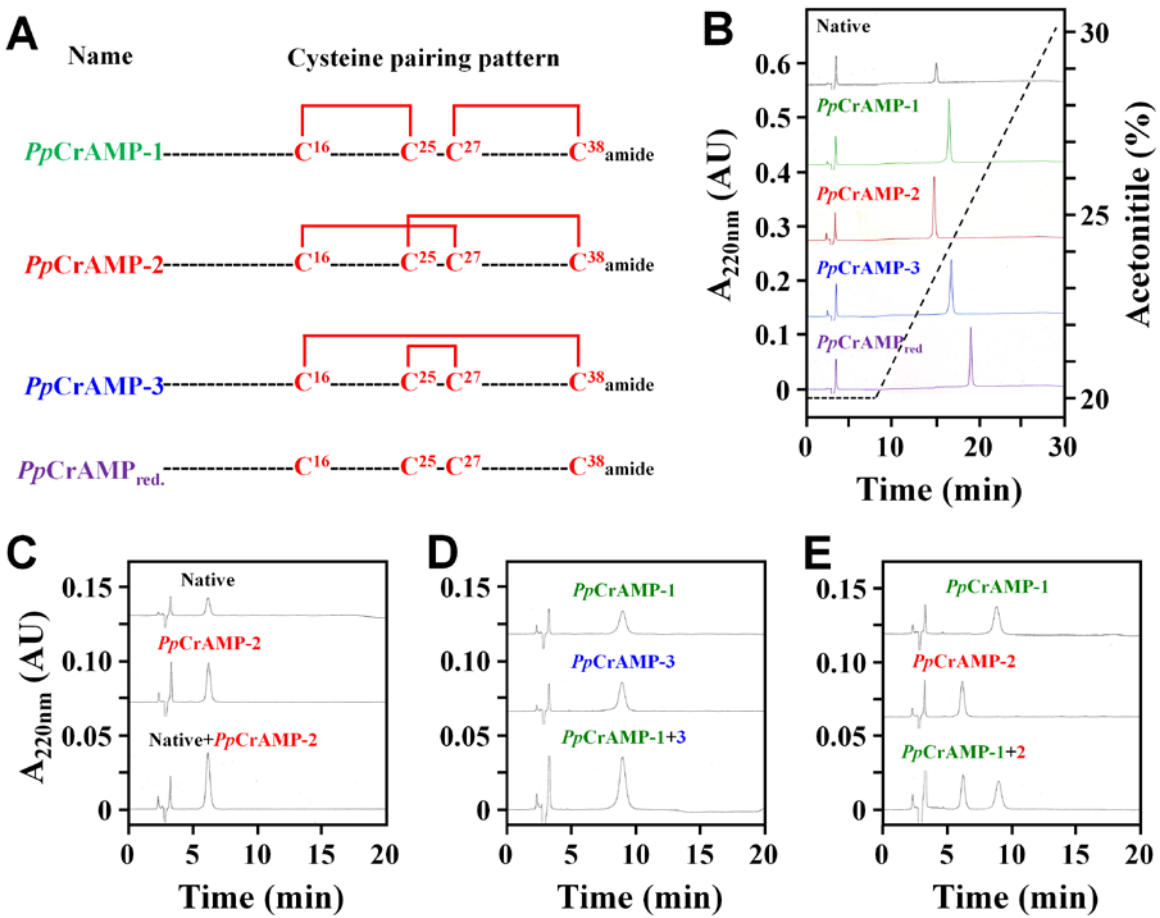
684

685 Figure 2.



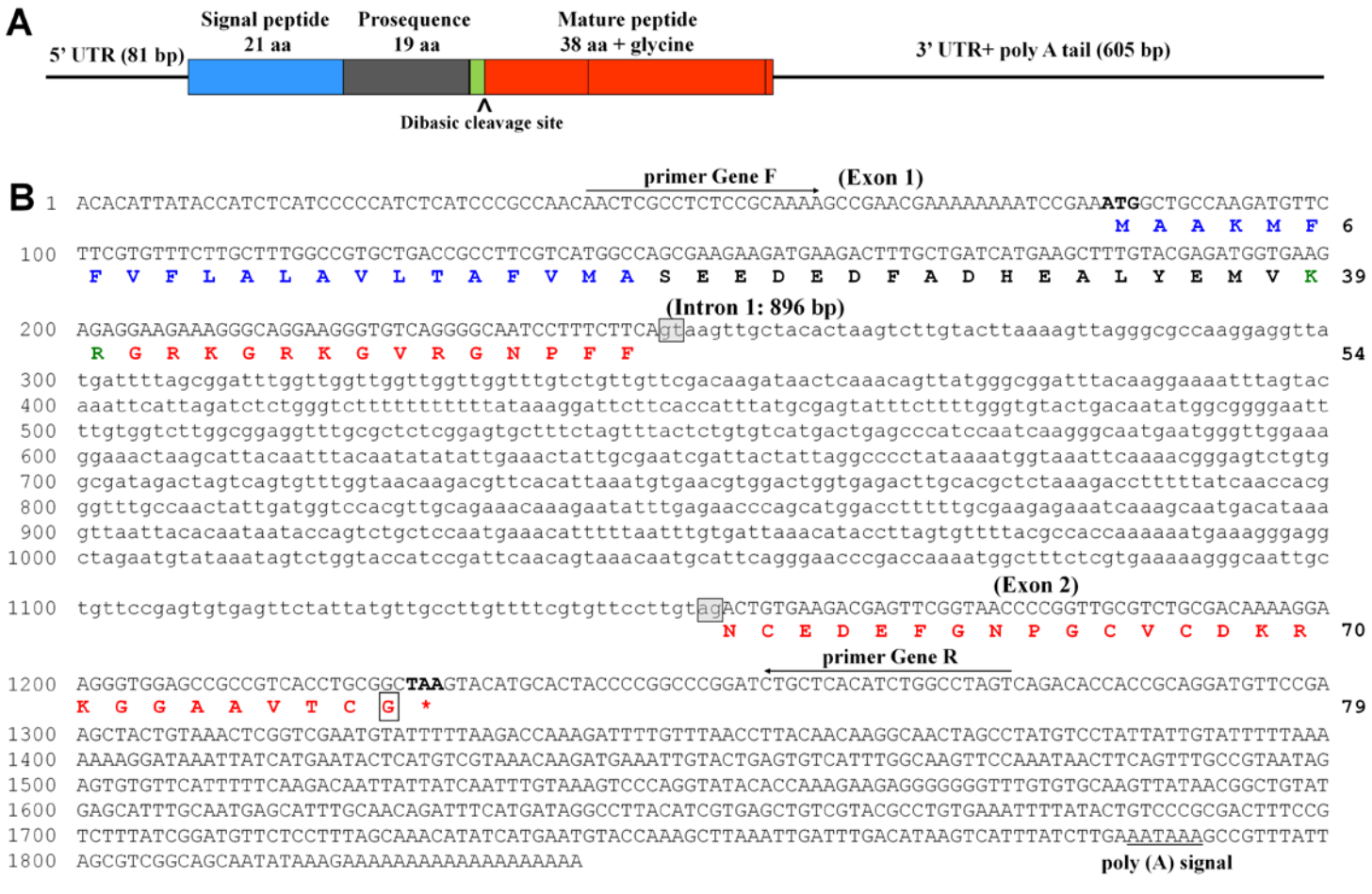
686

687 Figure 3.

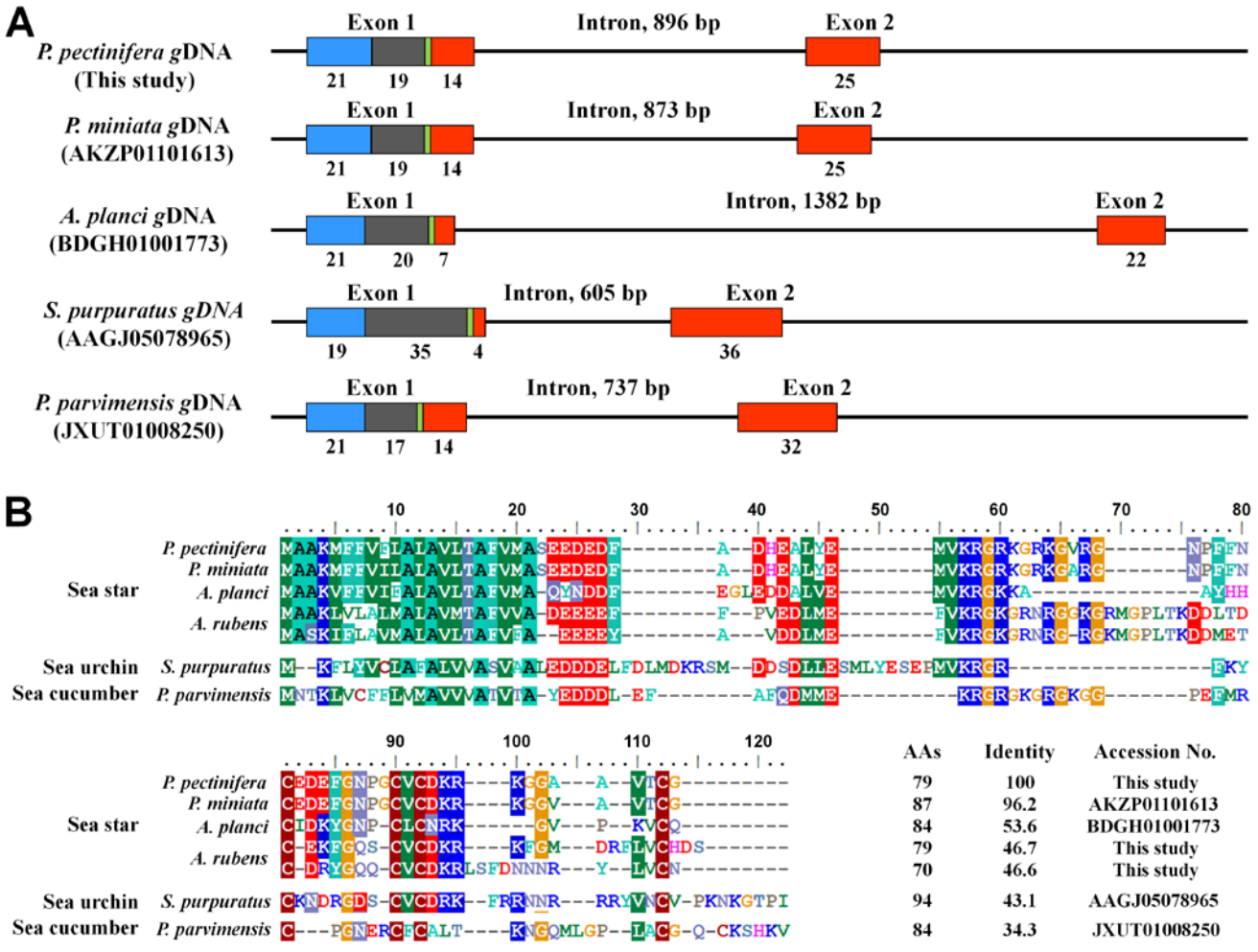


688

Figure 4.

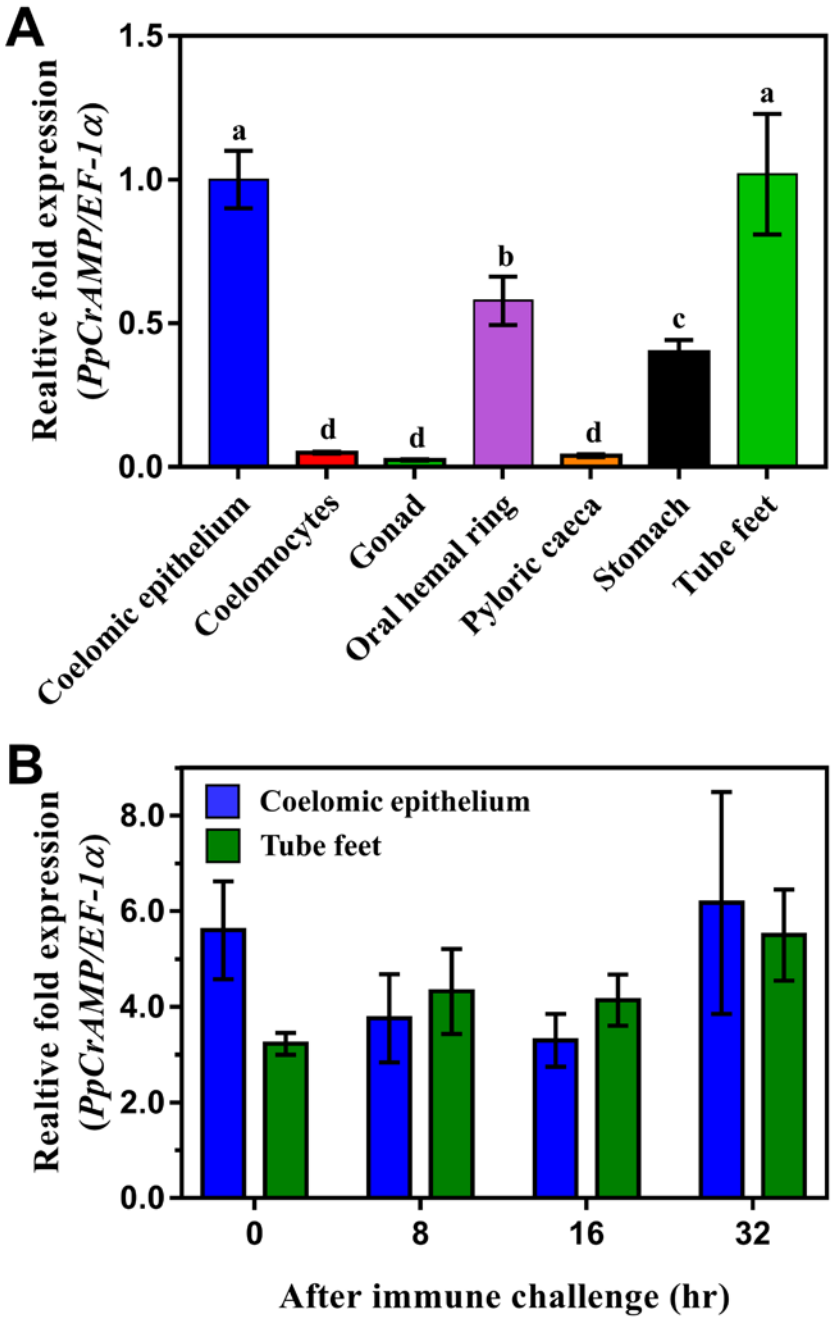

A
Signal peptide
21 aa
Prosequence
19 aa
Mature peptide
38 aa + glycine
5' UTR (81 bp)
3' UTR+ poly A tail (605 bp)
Dibasic cleavage site
B
primer Gene F
(Exon 1)
ACACATTATACCATCTCATCCCCCATCTCATCCCGCCAACAACTCGCCTCTCCGCAAAAGCCGAACGAAAAAAAATCCGAAATGGCTGCCAAGATGTTC
M A A K M F 6
TTCGTGTTTCTTGCTTTGGCCGTGCTGACCGCCTTCGTCATGGCCAGCGAAGAAGATGAAGACTTTGCTGATCATGAAGCTTTGTACGAGATGGTGAAG
F V F L A L A V L T A F V M A S E E D E D F A D H E A L Y E M V K 39
(Intron 1: 896 bp)
AGAGGAAGAAAGGGCAGGAAGGGTGTCAGGGGCAATCCTTTCTTCAgtaagttgctacactaagtcttgtacttaaaagttagggcgccaaggaggtta
R G R K G R K G V R G N P F F 54
tgattttagcggatttggttggttggttggttttgtctgttgttcgacaagataactcaaacagttatgggcggatttacaaggaaaatttagtac
aaattcattagatctctgggtctttttttttttataaaggattcttcaccatttatgcgagtatttcttttgggtgtactgacaatatggcggggaatt
ttgtggtcttggcggaggtttgcgctctcggagtgctttctagtttactctgtgtcatgactgagcccatccaatcaagggcaatgaatgggttggaaa
ggaaactaagcattacaatttacaatatatattgaaactattgcgaatcgattactattaggcccctataaaatggtaaattcaaaacgggagtctgtg
gcgatagactagtcagtgtttggtaacaagacgttcacattaaatgtgaacgtggactggtgagacttgcacgctctaaagacctttttatcaaccacg
ggtttgccaactattgatggtccacgttgcagaaacaaagaatatttgagaacccagcatggacctttttgcgaagagaaatcaaagcaatgacataaa
gttaattacacaataataccagtctgctccaatgaaacatttttaatttgtgattaaacataccttagtgttttacgccaccaaaaaatgaaagggagg
ctagaatgtataaatagtctggtaccatccgattcaacagtaaacaatgcattcagggaacccgaccaaaatggctttctcgtgaaaaagggcaattgc
(Exon 2)
tgttccgagtgtgagttctattatgttgccttgttttcgtgttccttgtagACTGTGAAGACGAGTTCGGTAACCCCGGTTGCGTCTGCGACAAAAGGA
N C E D E F G N P G C V C D K R 70
primer Gene R
AGGGTGGAGCCGCCGTCACCTGCGGCTAAGTACATGCACTACCCCGGCCCGGATCTGCTCACATCTGGCCTAGTCAGACACCACCGCAGGATGTTCCGA
K G G A A V T C G * 79
AGCTACTGTAAACTCGGTCGAATGTATTTTTAAGACCAAAGATTTTGTTTAACCTTACAACAAGGCAACTAGCCTATGTCCTATTATTGTATTTTTAAA
AAAAGGATAAATTATCATGAATACTCATGTCGTAAACAAGATGAAATTGTACTGAGTGTCATTTGGCAAGTTCCAAATAACTTCAGTTTGCCGTAATAG
AGTGTGTTCATTTTTCAAGACAATTATTATCAATTTGTAAAGTCCCAGGTATACACCAAAGAAGAGGGGGGGTTTGTGTGCAAGTTATAACGGCTGTAT
GAGCATTTGCAATGAGCATTTGCAACAGATTTCATGATAGGCCTTACATCGTGAGCTGTCGTACGCCTGTGAAATTTTATACTGTCCCGCGACTTTCCG
TCTTTATCGGATGTTCTCCTTTAGCAAACATATCATGAATGTACCAAAGCTTAAATTGATTTGACATAAGTCATTTATCTTGAAATAAAGCCGTTTATT
AGCGTCGGCAGCAATATAAAGAAAAAAAAAAAAAAAAAA
poly (A) signal

692 Figure 5.



693

694 Figure 6.



695

Figure 7.

Species	Peptide	Sequence	AAs	Cysteine connectivity
P. pectinifera	*Pp*CrAMP	GRKGRKGVRGNPFFNCEDEFGNPGCVCDKRKGGAAVTCa	38	C1-C3 and C2-C4
	Consensus secondary structure prediction	ccccccccccccccccccccccccccccccccccccccc		
S. scrofa	Protegrin-1	RGGRLCYCRRRFCVCVGRa	18	C1-C4 and C2-C3
A. australis	Androctonin	RSVCRQIKICRRRGGCYYKCTNRPY	25	
A. gomesiana	Gomesin	pQCRRLCYKQRCVTYCRGRa	18	
L. polyphemus	Polyphemusin-I	RRWCFRVCYRGFCYRKCRa	18	
T. tridentatus	Tachyplesin-I	KWCFRVCYRGICYRRCRa	17	

698 **Table 1.** Antimicrobial activity against various microbial strains using synthetic *Pp*CrAMPs, including
699 a reduced linear peptide and peptides with two disulfide bonds in three different configurations

Microbe	[a]Minimal effective concentration (μg/ml)			
	***Pp*CrAMP-1**	***Pp*CrAMP-2**	***Pp*CrAMP-3**	***Pp*CrAMP$_{red.}$**
Gram-positive				
Bacillus subtilis KCTC1021	33.8	22.9	38.3	42.3
Staphylococcus aureus KCTC1621	32.1	15.6	41.2	91.2
Micrococcus luteus KCTC1071	>250	153	>250	82.9
Gram-negative				
Escherichia coli D31	[b]ND	ND	ND	ND
Salmonella enterica ATCC13311	8.0	4.5	8.1	8.4
Shigella flexneri KCTC2517	29.8	12.2	31.4	24.2
Marine bacterium (gram-positive)				
Streptococcus iniae FP5229	ND	ND	ND	ND
Marine bacteria (gram-negative)				
Aeromonas hydrophila KCTC2358	ND	107.2	ND	ND
Edwardsiella tarda NUF251	ND	ND	ND	ND
Vibrio parahaemolyticus KCCM41664	ND	>250	ND	ND
Fungus				
Candida albicans KCTC9765	ND	ND	ND	ND

700 [a]Antimicrobial assay were performed in triplicates and the results were averaged.

701 [b]ND means not detected in the range of the concentrations tested up to 250 μg/ml of peptides

Supplementary Table S1. Designations and nucleotide sequences of the primers used in this study

Primers	Nucleotide sequence (5'→3')	Usage
GSP-F1	GGTGTCAGGGGCAATCCTTT	cDNA cloning
GSP-F2	CAACTGTGAAGACGAGTTCGG	
GSP-R	GCATGTACTTAGCCGCAGG	
Gene F	AACTCGCCTCTCCGCAAAA	Gene cloning
Gene R	ACTAGGCCAGATGTGAGCAG	
PpCrAMP qPCR-F	GGTGTCAGGGGCAATCCTTT	RT-qPCR
PpCrAMP qPCR-R	GGCTCCACCCTTCCTTTTGT	
EF1α qPCR-F	TCAACGACTACCAGCCCCTA	
EF1α qPCR-R	TTCTTGCTAGCCTTCTGGGC	

Highlights

- A novel cysteine-rich AMP (*Pp*CrAMP) is identified from the starfish *Patiria pectinifera*.
- *Pp*CrAMP adopts two disulfide bonds with Cys^{16}-C^{27} and Cys^{25}-Cys^{38} pairings.
- *PpCrAMP* transcripts are highly expressed in the tube feet and the coelomic epithelium.
- *PpCrAMP* gene contains an intron.
- *Pp*CrAMP exhibits antimicrobial activity to different bacteria.